(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
19 September 2002 (19.09.2002)

PCT

(10) International Publication Number
WO 02/073954 A1

(51) International Patent Classification[7]: H04N 5/262

(21) International Application Number: PCT/JP02/02298

(22) International Filing Date: 12 March 2002 (12.03.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2001-071121 13 March 2001 (13.03.2001) JP

(71) Applicant *(for all designated States except US)*: **CANON KABUSHIKI KAISHA** [JP/JP]; 3-30-2, Shimomaruko, Ohta-ku, Tokyo 146-8501 (JP).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SHIMOYAMA, Tomohiko** [JP/JP]; 3-7-9-1405, Hikarigaoka, Nerima-ku, Tokyo 179-0072 (JP). **MORITA, Kenji** [JP/JP]; 4-8-27-104, Tsunashimahigashi, Kohoku-ku, Yokohama-shi, Kanagawa 223-0052 (JP). **YONEZAWA, Hiroki** [JP/JP]; 10-20-202, Ichibahigashinakacho, Tsurumi-ku, Yokohama-shi, Kanagawa 230-0022 (JP).

(74) Agent: **OHTSUKA, Yasunori**; 7th FL., SHUWA KIOICHO PARK BLDG., 3-6, KIOICHO, CHIYODA-KU, Tokyo 102-0094 (JP).

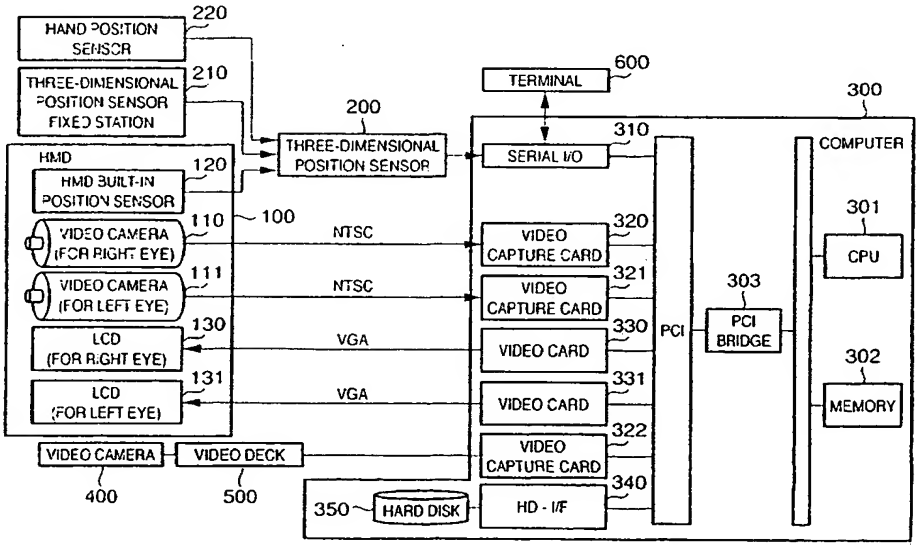
(81) Designated State *(national)*: US.

(84) Designated States *(regional)*: European patent (DE, FR, GB, IT, NL).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: IMAGE COMPOSITION APPARATUS AND METHOD



(57) Abstract: ABSTRACT An image composition system includes an HMD (100) having a right-eye camera (110), a left-eye camera (111), a right-eye LCD (130) and a left-eye LCD (131) for displaying a real image, and the like, and an information processing apparatus (300) for generating another image different from the real image. A composite image obtained by superimposing the other image generated by the information processing apparatus (300) on the real image captured by the right-eye camera (110) and left-eye camera (111) is displayed on the right-eye LCD (130) and left-eye LCD (131). The display region of the other image is determined based on the position and posture of the head of the user who wears the HMD (100). The user can observe the other image superimposed on the real image at an appropriate position while wearing the HMD on his or her head.

WO 02/073954 A1

DESCRIPTION

IMAGE COMPOSITION APPARATUS AND METHOD

TECHNICAL FIELD

The present invention relates to an image composition apparatus that displays a real image superimposed with another image on a display unit of display means to be worn on a head.

BACKGROUND ART

Conventionally, upon shooting movie or television program scenes, a performer acts according to memorized script contents. After shooting one scene, a director gives directions about that scene to the performer, the performer confirms the directions while observing playback of a video of himself or herself, and shooting progresses while reflecting those directions in action. In such process, shooting is made.

However, it is a heavy burden for a performer to memorize the script contents. Since the director gives directions after shooting one scene, the performer cannot receive fine directions from the director during action. Also, the performer cannot see a video of himself or herself, i.e., how he or she is acting, during shooting.

DISCLOSURE OF INVENTION

description, serve to explain the principles of the invention.

Fig. 1 shows a use example of an image composition apparatus according to an embodiment of the present invention;

Fig. 2 is a block diagram showing the use example of Fig. 1;

Fig. 3A is a perspective view of an HMD (Head Mount Display) in Figs. 1 and 2 when viewed from the front side;

Fig. 3B is a perspective view of the HMD (Head Mount Display) in Figs. 1 and 2 when viewed from the rear side;

Fig. 4 shows the generation processes of a video to be superimposed on the HMD;

Fig. 5 is a diagram showing the configuration of programs which run on an information processing apparatus 300 shown in Fig. 2;

Fig. 6 is a flow chart showing the process of an HMD display thread 1000 shown in Fig. 5;

Fig. 7 is a flow chart showing the process for determining a display position in step S103 in Fig. 6;

Fig. 8 is a flow chart showing the process of a terminal management thread 2000 in Fig. 5;

Fig. 9 is a flow chart showing the process of a script management thread 3000 in Fig. 5; and

Fig. 10 is a state transition chart showing transition of the state of a gesture recognition thread 4000.

combination of an optical see-through HMD and video camera in place of the video see-through HMD.

The right-eye camera 110 of the HMD 100 is connected to a video capture card 320 of the information processing apparatus 300, and the left-eye camera 111 is connected to a video capture card 321 of the information processing apparatus 300. The right-eye LCD 130 is connected to a video card 330 of the information processing apparatus 300, and the left-eye LCD 131 is connected to a video card 331 of the information processing apparatus 300. The LCDs 130 and 131 display a composite video of those actually captured by the left-eye camera 111 and right-eye camera 110, and, for example, script data or a video from the video camera 400 (Fig. 4).

The video to be displayed on the HMD 100 is generated by the information processing apparatus 300. The information processing apparatus 300 comprises a CPU 301, memory 302, PCI bridge 303, hard disk I/F 340, hard disk 350, and the like in addition to a serial I/O 310, the video capture cards 320, 321, and 322, and video cards 330 and 331 mentioned above.

The three-dimensional position sensor 200 comprises the three-dimensional position sensor fixed station 210 and the three-dimensional sensor mobile station 120 which is built in the HMD. The three-dimensional position sensor 200 measures the relative position between the three-dimensional position sensor fixed station 210 and

This embodiment uses a magnetic position sensor which is separated into a fixed station and position sensor, but may use a position sensor using a gyro as long as it can measure the position of the HMD. The terminal 600 is used to input instructions from a staff member or to input shooting start and stop instructions.

The configuration of programs which run on the information processing apparatus 300 will be explained below. In the following description, assume that a performer wears the HMD 100 in rehearsal upon shooting a movie or television program.

Fig. 5 shows the configuration of programs which run on the information processing apparatus 300 in Fig. 2. The programs include an HMD display thread 1000, terminal management thread 2000, script management thread 3000, and gesture recognition thread 4000. Data are exchanged among the threads via an instruction buffer 2001, script buffer 3002, and display mode flag 4001.

The HMD display thread 1000 displays videos captured by the right-eye camera 110 and left-eye camera 111 on the LCDs 130 and 131. In this case, the thread 1000 superimposes an instruction from a staff member written in the instruction buffer 2001 or script data written in the script buffer 3002. Also, an image taken by the television camera 400 is also superimposed.

determination method of the information display position will be explained later.

After the information display position is determined, a video from the video camera 400, script data 350, and an instruction from a staff member via the terminal 600 are captured (step S104), and the captured information is written in a video buffer area corresponding to the determined display position (step S105). Since the video from the HMD 100 has already been written in the video buffer, information is superimposed on that video.

Upon completion of rendering, the contents of the video buffer are transferred to a frame buffer on the video board 330 to display (render) the contents (video) of the video buffer on the LCD 130 in the HMD 100 (step S106).

The determination method of the information display position in step S103 will be described below.

Fig. 7 is a flow chart of the process for determining the information display position in step S103 in Fig. 6.

The position of the HMD 100 is acquired from the three-dimensional sensor main body 200 (step S200). The information processing apparatus 300 generates and sets a modeling conversion matrix on the basis of the position acquired in step S200, the coordinate position of the HMD 100, and parameters such as the focal length of the camera and the like, which are measured in advance, so as to obtain an image from the viewpoint of the HMD 100 (step S201). Note that the "position and posture of the viewpoint of an

video camera 400 and information from the script data 350 are determined to be the center of the four corners of the table, and the video and the information are rendered, as indicated by a video 4c. The videos 4a and 4c are composited to obtain a video 4d, which is observed by the person who wears the HMD 100. Note that Fig. 4 typically illustrates the respective videos, and some of the video contents, composite positions, and the like are not accurate.

In this embodiment, as described previously, since the information is superimposed on the table, the coordinate position of which is known, the field of view of the performer can be prevented from being intercepted. Display of information on the table does not limit the present invention. For example, information may be superimposed on a portion of a wall, the coordinate position of which is known. Furthermore, the position of the display can be dynamically changed. For example, the position of the display may be changed from a desk to a wall.

The terminal management thread 2000 in Fig. 5 mainly processes an input from the terminal 600, and writes an instruction from a staff member to the performer via the terminal 600 in the instruction buffer 2001. At the same time, the terminal management thread 2000 informs the script management thread 3000 of staff member's operations.

The process of the terminal management thread 2000 will be described below.

The script data is stored as a sequence of sets of time stamps and character strings to be displayed at those timings.

The process of the script management thread 3000 will be described below.

Fig. 9 is a flow chart showing the process of the script management thread 3000 in Fig. 5.

As an initial setup process, an internal shooting clock 3001 is reset to zero, and a pointer of script data is returned to the head of the script (step S400).

Data (next script data to be displayed) pointed by the pointer of the script data is loaded from the hard disk 350 (step S401). The control waits until the shooting clock 3001 becomes the same as the time stamp of the script data pointed by the pointer (step S402).

The script data pointed by the pointer is written in the script buffer (step S403). The pointer of the script data is advanced (step S404), and the flow returns to step S401 to repeat the aforementioned steps.

The gesture recognition thread 4000 in Fig. 5 recognizes a gesture (hand position and posture) of the performer on the basis of the position of the hand position sensor 220 obtained via the three-dimensional position sensor main body 200. Every time a gesture is recognized, the display mode flag 4001 is turned on/off.

In this embodiment, as a gesture for turning on/off display, an action for moving the hand up and down three times for a second is selected.

When a downward acceleration is detected in the "upward acceleration stop state S502", the gesture recognition thread 4000 transits to a "downward acceleration state S503". When the acceleration has stopped, the gesture recognition thread 4000 transits to a "downward acceleration stop state S504". When a downward acceleration is detected again within 0.1 sec after transition, the gesture recognition thread 4000 returns to the "downward acceleration state S503".

When an upward acceleration is detected in the "downward acceleration stop state S504", the gesture recognition thread 4000 transits to the "upward acceleration state S501". At this time, the internal counter is incremented. This corresponds to a case wherein the hand is moved downward after upward movement.

When no acceleration is detected within 0.1 sec after transition to the "downward acceleration stop state S504", the gesture recognition thread 4000 transits to the "standby state S500". If counter = 3, it is determined that the gesture is complete, and an event is generated to invert the value (TRUE/FALSE) of the display mode flag 4001.

The gesture recognition thread 4000 executes the process according to the aforementioned state transition chart to detect an event. In the above description, a full-superimpose display ON/OFF instruction is issued by a gesture. However, for example, display of script data, instruction data, and video camera image may be

In this case, the program code itself read out from the storage medium implements the functions of the above-mentioned embodiments, and the storage medium which stores the program code constitutes the present invention.

As the storage medium for supplying the program code, for example, a floppy disk, hard disk, optical disk, magneto-optical disk, CD-ROM, CD-R, magnetic tape, nonvolatile memory card, ROM, and the like may be used.

The functions of the above-mentioned embodiments may be implemented not only by executing the readout program code by the computer but also by some or all of actual processing operations executed by an OS (operating system) running on the computer on the basis of an instruction of the program code.

Furthermore, the functions of the above-mentioned embodiments may be implemented by some or all of actual processing operations executed by a CPU or the like arranged in a function extension board or a function extension unit, which is inserted in or connected to the computer, after the program code read out from the storage medium is written in a memory of the extension board or unit.

As described in detail above, according to the image composition apparatus, since another image is displayed on a display unit that displays a real image, the other image can be superimposed on the real image. Hence, the user who wears display means on the head can observe the other image superimposed on the real image.

CLAIMS

1. An image composition system for compositing a real image in a line-of-sight direction of a user with another image, comprising:

a display unit which is wearable on a head of the user, and displays a composite image;

a position sensor for detecting the line-of-sight direction of the user, and outputting line-of-sight information;

a determination unit for determining a display region where the other image is to be displayed, in accordance with the line-of-sight information; and

a composition unit for compositing the other image on the determined display region,

wherein the other image is used to display information that helps operations of the user.

2. The system according to claim 1, wherein said display unit has an optical see-through structure, and the user can observe a real space via said display unit.

3. The system according to claim 1, further comprising:

a first image taking device for obtaining a video of a real space observed from a viewpoint of the user, and

wherein said composition unit displays the video obtained by said first image taking device on said display unit, and superimposes the other image on the display region determined by said determination unit.

9. The system according to claim 8, wherein said composition unit switches contents of the other image to be displayed on the display region in response to a predetermined action detected by said gesture detection unit.

10. The system according to claim 1, wherein the information that helps the operations of the user is dialog information.

11. The system according to claim 1, wherein the information that helps the operations of the user is an image obtained by taking an image of an action of the user.

12. An information processing method of displaying a composite image of a real image in a line-of-sight direction of a user and another image on a display unit which is wearable on a head of the user, comprising the steps of:

detecting the line-of-sight direction of the user to acquire line-of-sight information;

determining a display region where the other image is to be displayed, in accordance with the line-of-sight information; and

compositing the other image on the determined display region,

wherein the other image is used to display information that helps operations of the user.

13. The method according to claim 12, wherein the other image information is a video obtained by image taking means

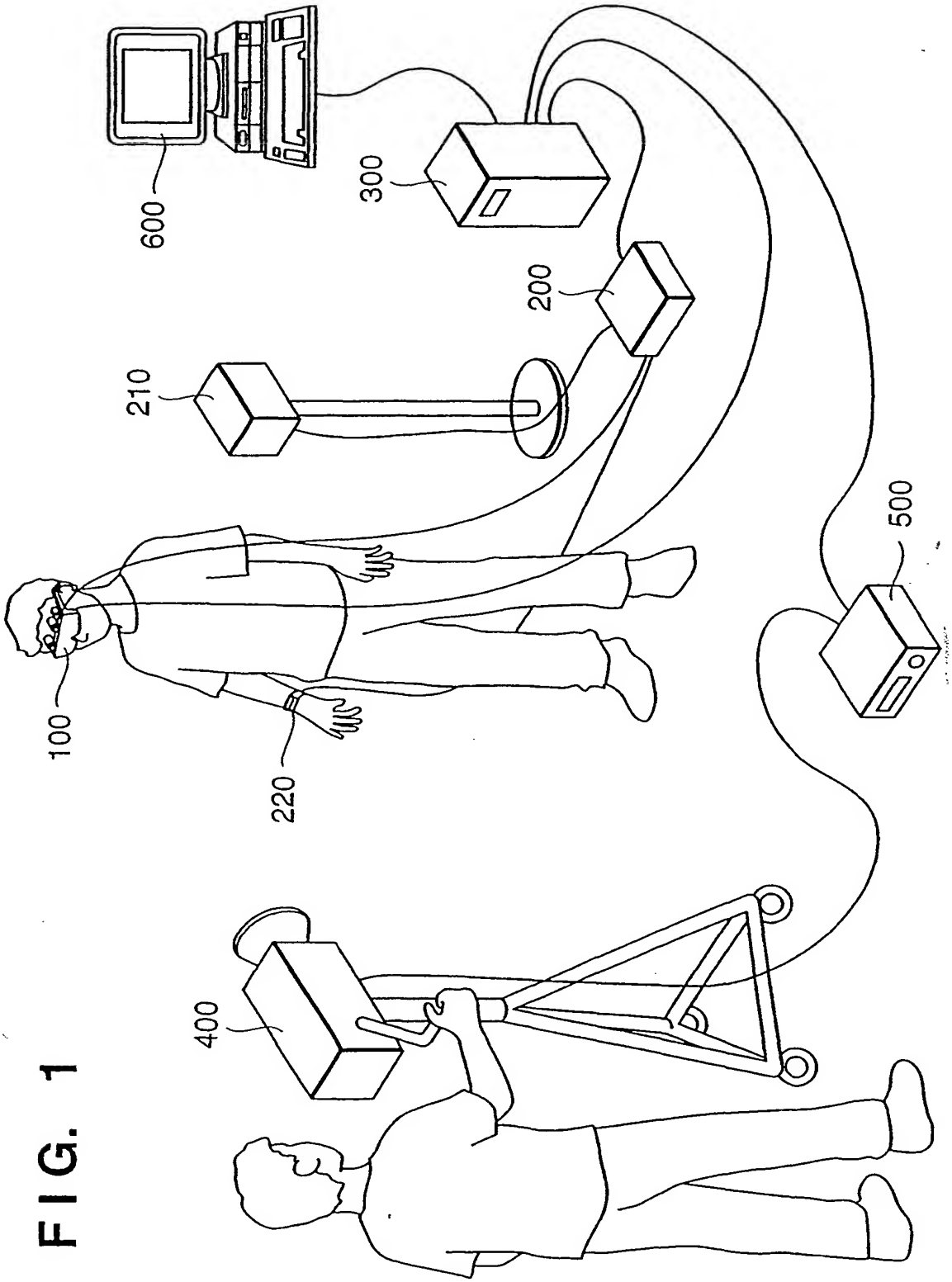


FIG. 1

FIG. 2

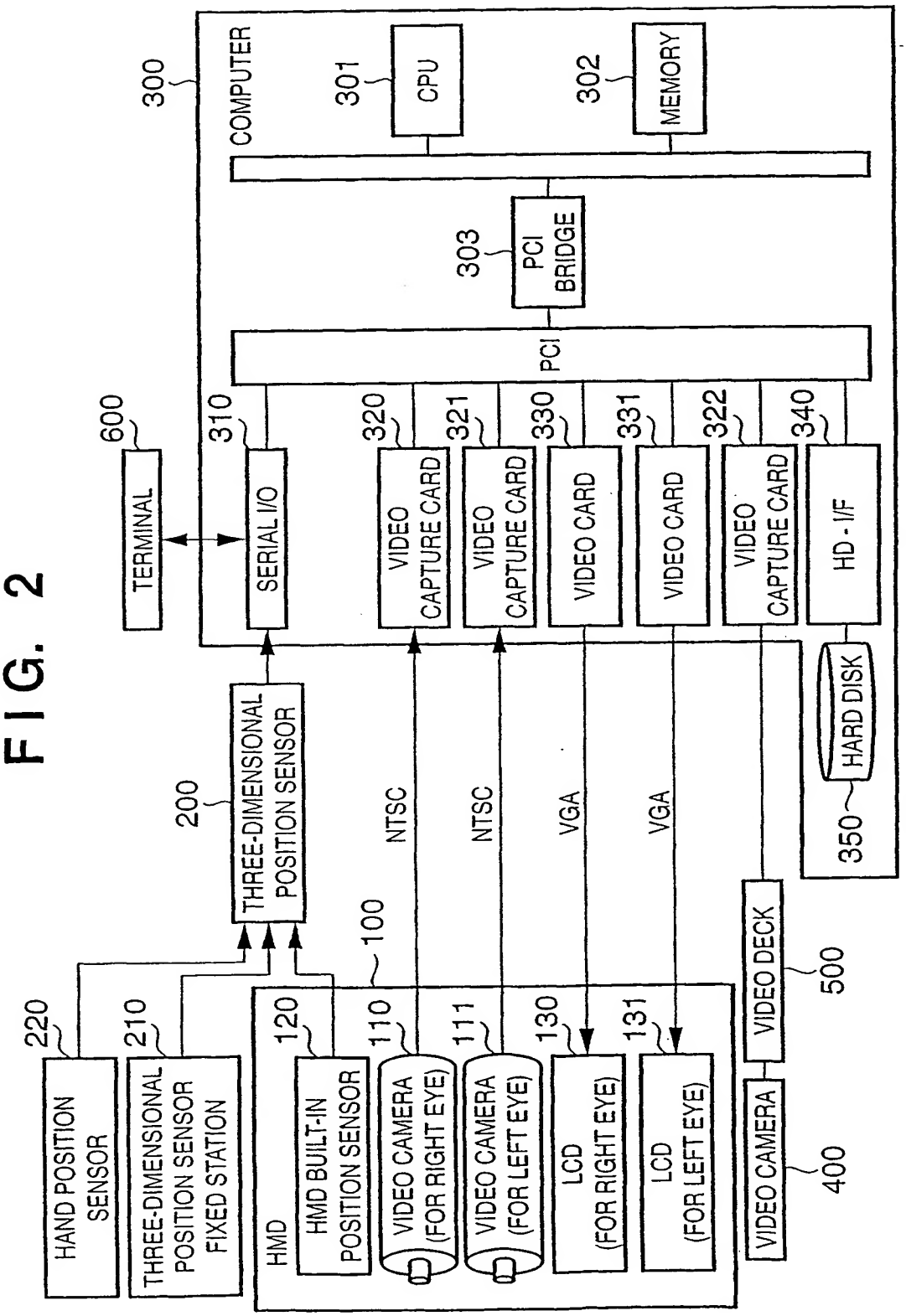

300
COMPUTER
301
CPU
302
MEMORY
303
PCI BRIDGE
PCI
600
TERMINAL
310
SERIAL I/O
320
VIDEO CAPTURE CARD
321
VIDEO CAPTURE CARD
330
VIDEO CARD
331
VIDEO CARD
322
VIDEO CAPTURE CARD
340
HD - I/F
HARD DISK
350
200
THREE-DIMENSIONAL POSITION SENSOR
NTSC
NTSC
VGA
VGA
100
220
HAND POSITION SENSOR
210
THREE-DIMENSIONAL POSITION SENSOR FIXED STATION
HMD
120
HMD BUILT-IN POSITION SENSOR
110
VIDEO CAMERA (FOR RIGHT EYE)
111
VIDEO CAMERA (FOR LEFT EYE)
130
LCD (FOR RIGHT EYE)
131
LCD (FOR LEFT EYE)
VIDEO DECK
500
VIDEO CAMERA
400

FIG. 3A

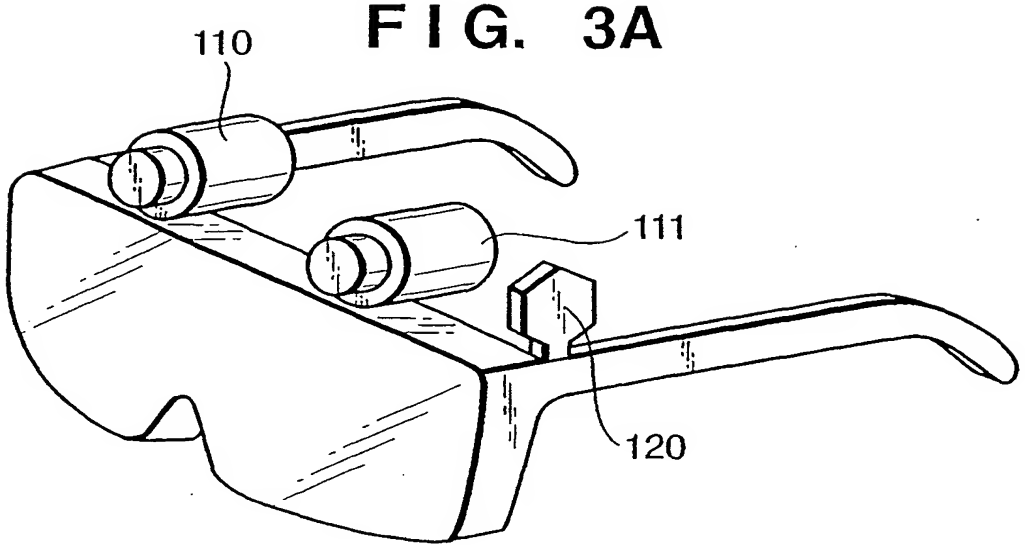


FIG. 3B

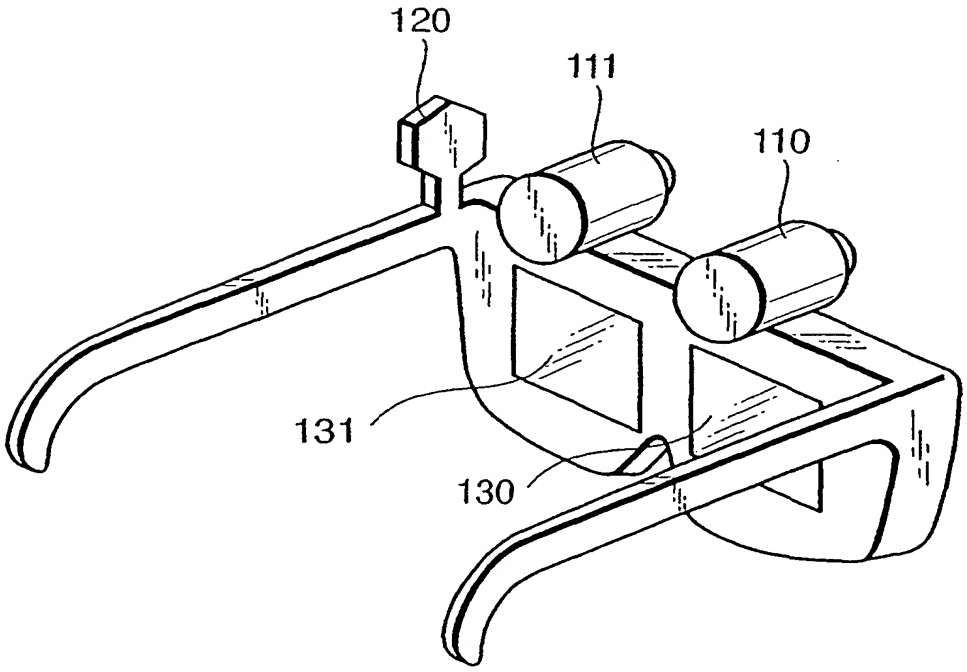


FIG. 4

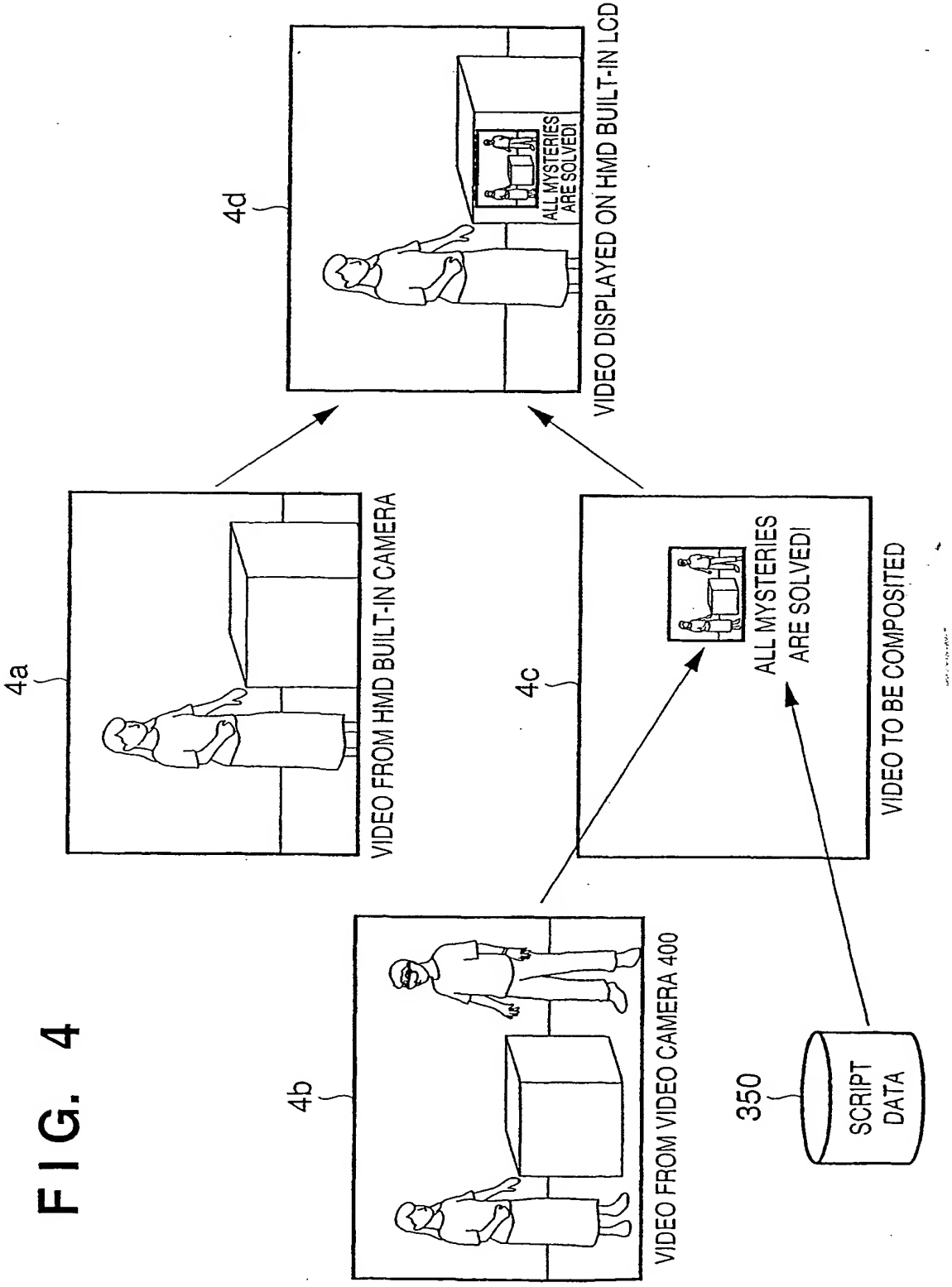

4d
ALL MYSTERIES ARE SOLVED!
VIDEO DISPLAYED ON HMD BUILT-IN LCD
4a
VIDEO FROM HMD BUILT-IN CAMERA
4c
ALL MYSTERIES ARE SOLVED!
VIDEO TO BE COMPOSITED
4b
VIDEO FROM VIDEO CAMERA 400
350
SCRIPT DATA

FIG. 5

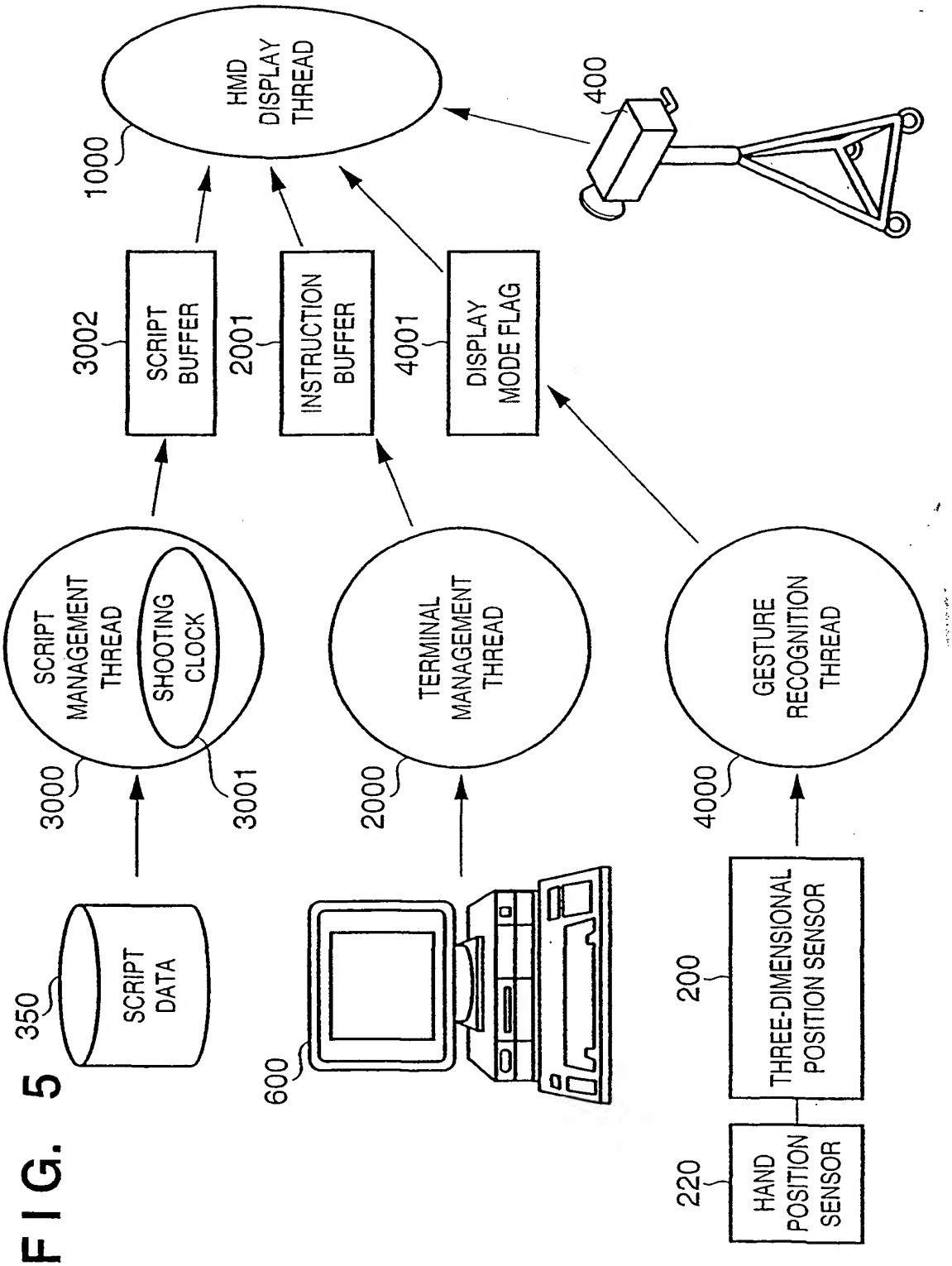

HMD DISPLAY THREAD
1000
400
SCRIPT BUFFER
3002
INSTRUCTION BUFFER
2001
DISPLAY MODE FLAG
4001
SCRIPT MANAGEMENT THREAD
SHOOTING CLOCK
3000
3001
TERMINAL MANAGEMENT THREAD
2000
GESTURE RECOGNITION THREAD
4000
SCRIPT DATA
350
600
THREE-DIMENSIONAL POSITION SENSOR
200
HAND POSITION SENSOR
220

FIG. 6

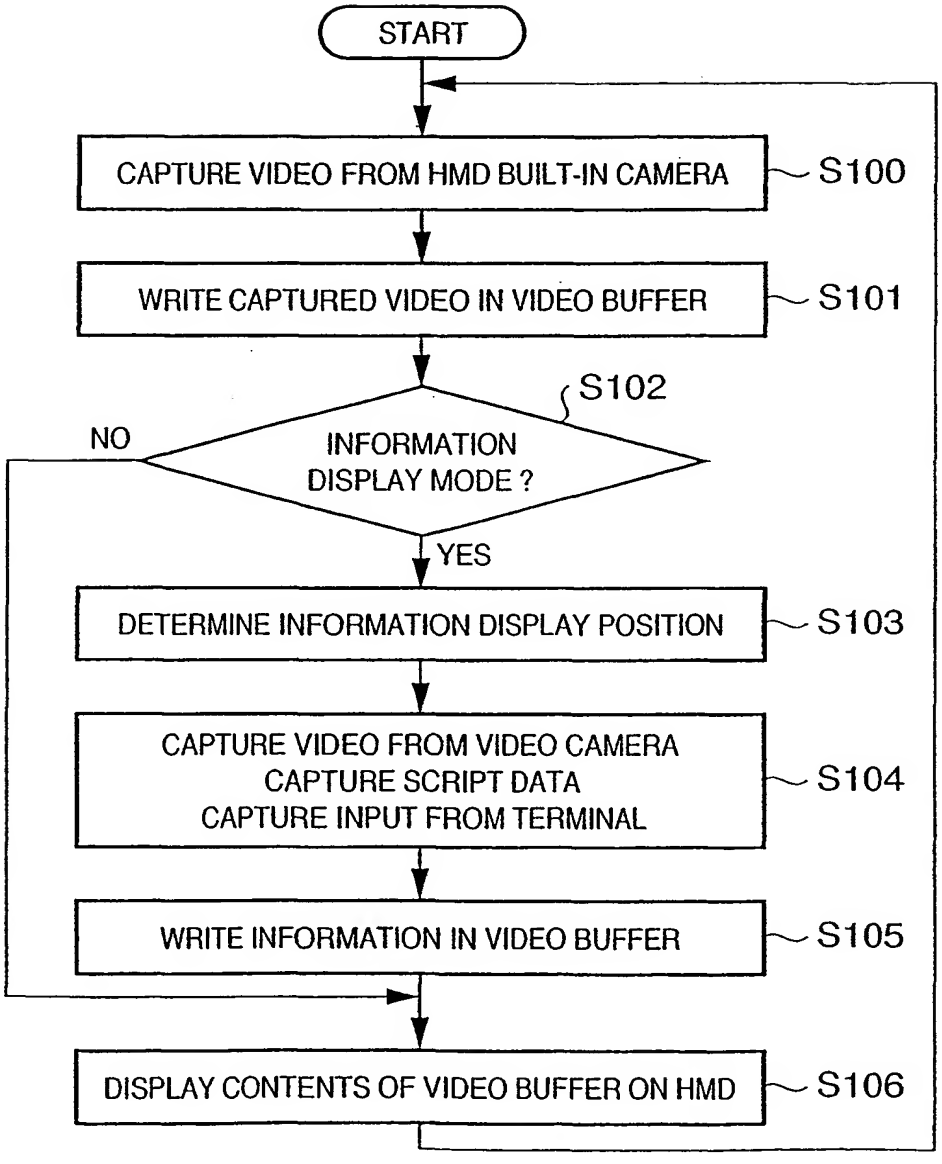

START
CAPTURE VIDEO FROM HMD BUILT-IN CAMERA
S100
WRITE CAPTURED VIDEO IN VIDEO BUFFER
S101
S102
INFORMATION DISPLAY MODE ?
NO
YES
DETERMINE INFORMATION DISPLAY POSITION
S103
CAPTURE VIDEO FROM VIDEO CAMERA
CAPTURE SCRIPT DATA
CAPTURE INPUT FROM TERMINAL
S104
WRITE INFORMATION IN VIDEO BUFFER
S105
DISPLAY CONTENTS OF VIDEO BUFFER ON HMD
S106

FIG. 7

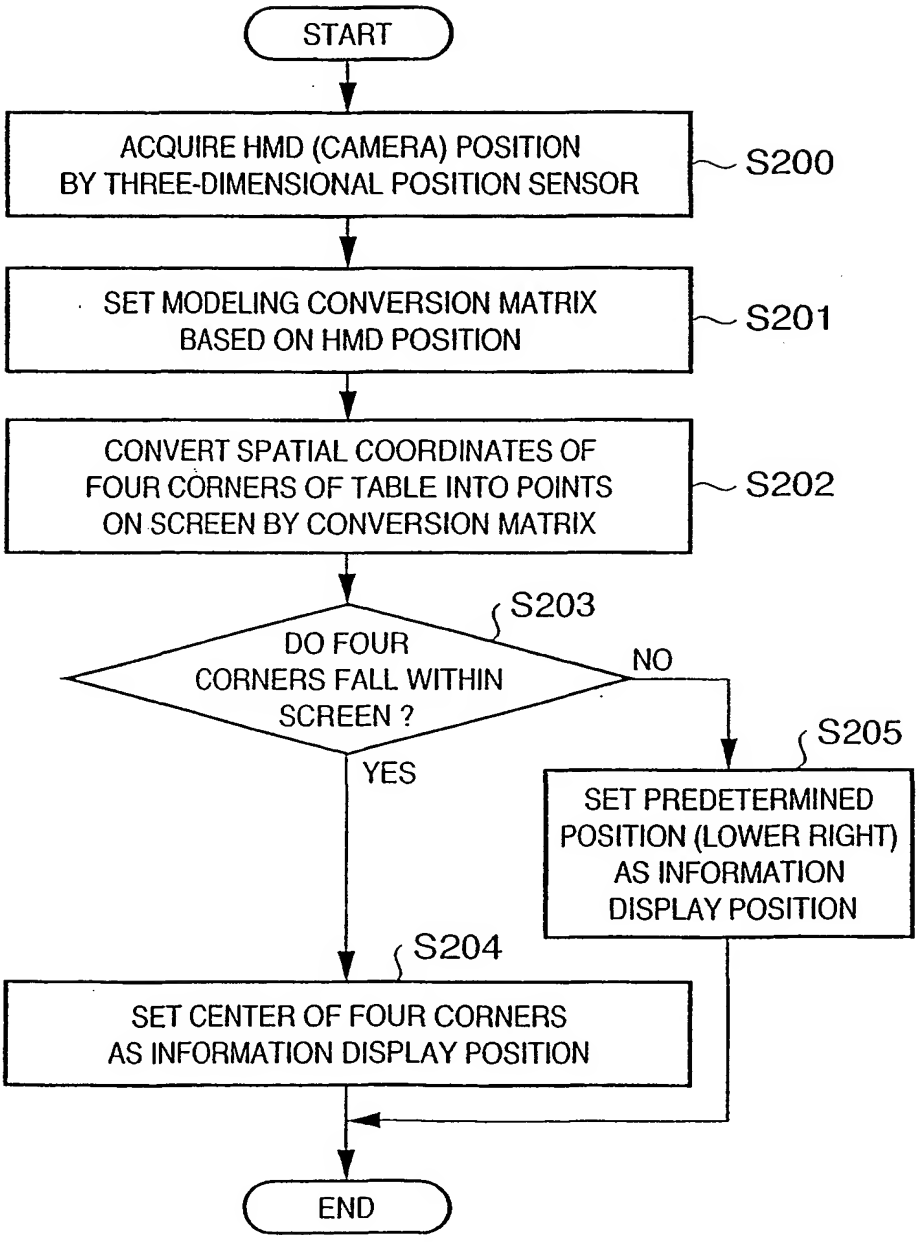

START
ACQUIRE HMD (CAMERA) POSITION BY THREE-DIMENSIONAL POSITION SENSOR
S200
SET MODELING CONVERSION MATRIX BASED ON HMD POSITION
S201
CONVERT SPATIAL COORDINATES OF FOUR CORNERS OF TABLE INTO POINTS ON SCREEN BY CONVERSION MATRIX
S202
S203
DO FOUR CORNERS FALL WITHIN SCREEN ?
NO
YES
S205
SET PREDETERMINED POSITION (LOWER RIGHT) AS INFORMATION DISPLAY POSITION
S204
SET CENTER OF FOUR CORNERS AS INFORMATION DISPLAY POSITION
END

FIG. 8

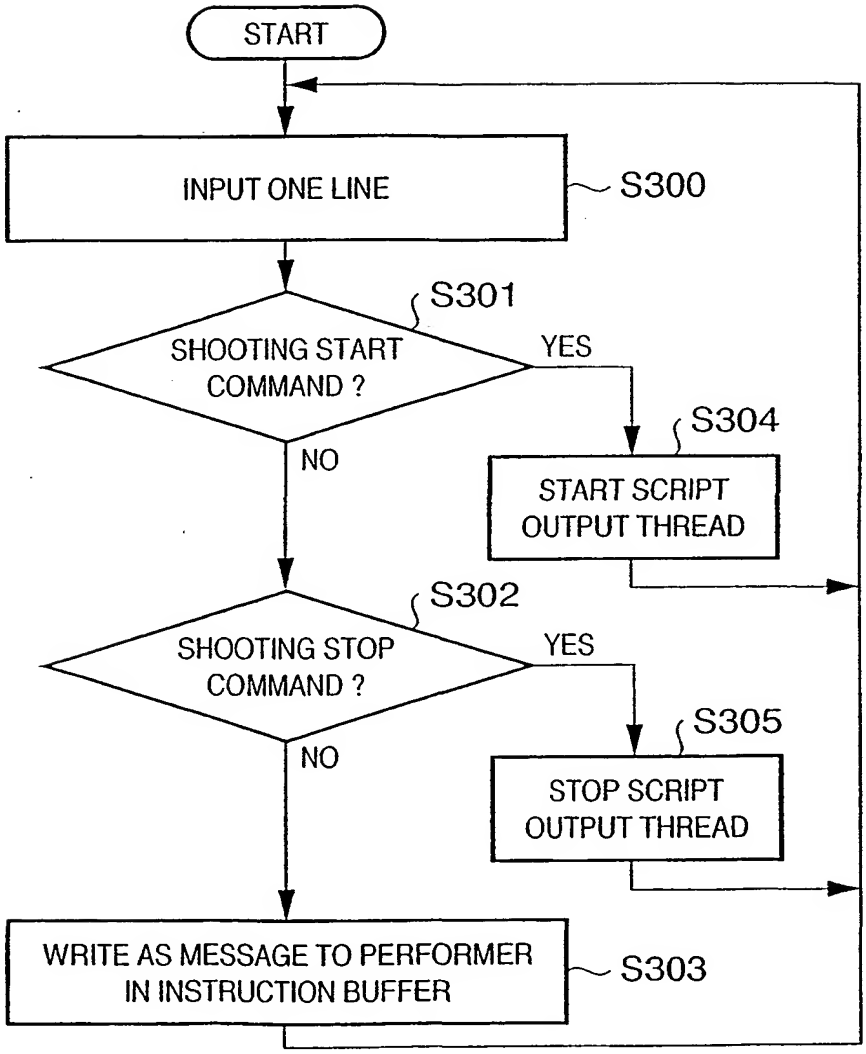

START
INPUT ONE LINE
S300
S301
SHOOTING START COMMAND ?
YES
NO
S304
START SCRIPT OUTPUT THREAD
S302
SHOOTING STOP COMMAND ?
YES
NO
S305
STOP SCRIPT OUTPUT THREAD
WRITE AS MESSAGE TO PERFORMER IN INSTRUCTION BUFFER
S303

FIG. 9

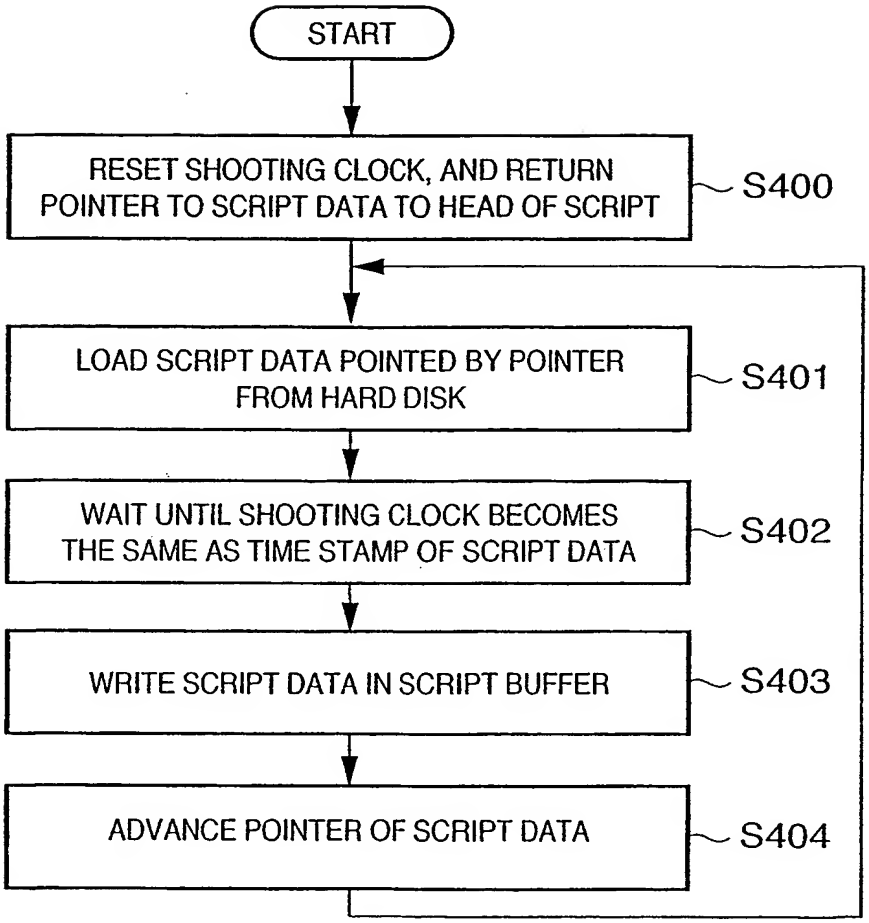
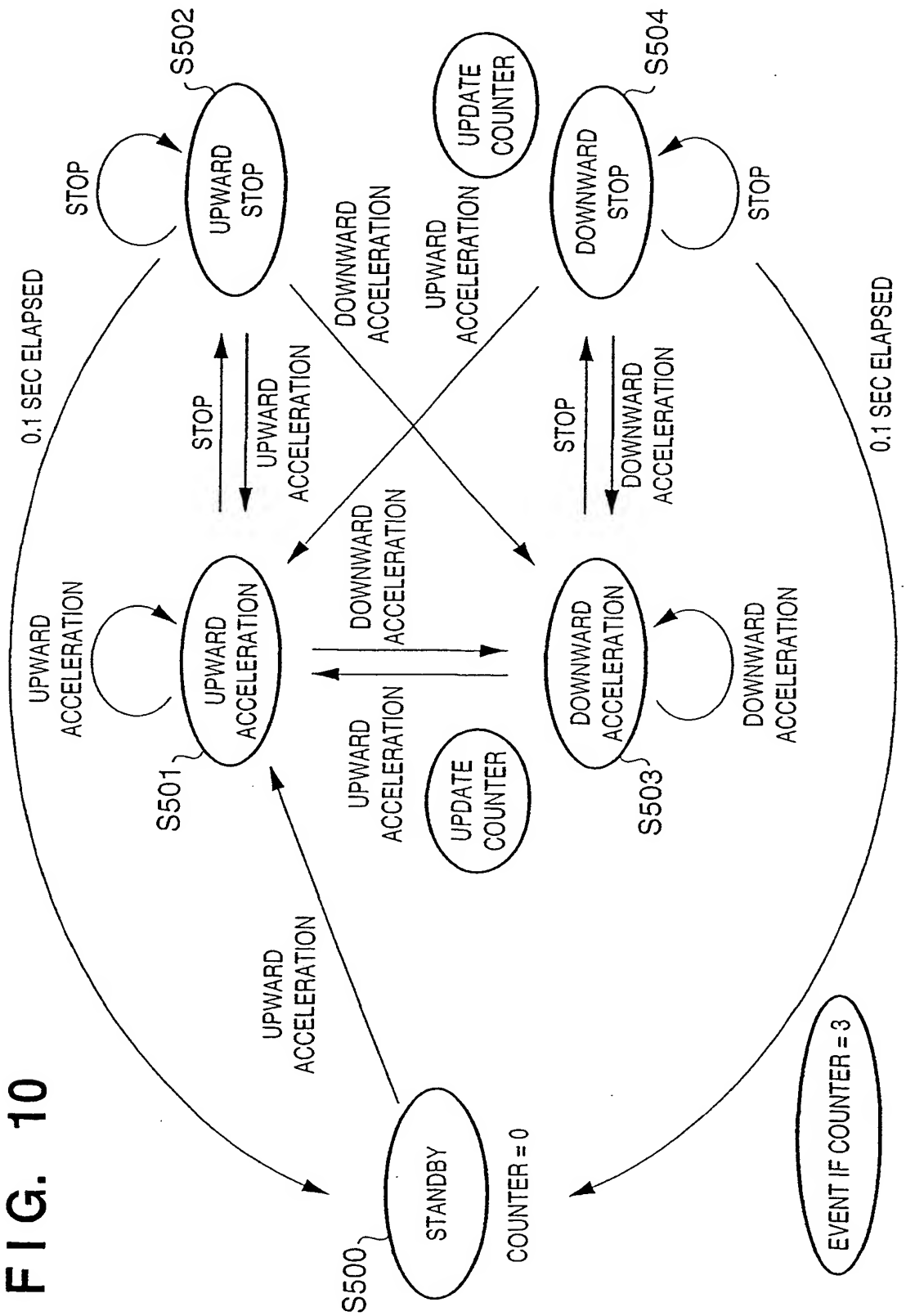

START
RESET SHOOTING CLOCK, AND RETURN POINTER TO SCRIPT DATA TO HEAD OF SCRIPT
S400
LOAD SCRIPT DATA POINTED BY POINTER FROM HARD DISK
S401
WAIT UNTIL SHOOTING CLOCK BECOMES THE SAME AS TIME STAMP OF SCRIPT DATA
S402
WRITE SCRIPT DATA IN SCRIPT BUFFER
S403
ADVANCE POINTER OF SCRIPT DATA
S404

FIG. 10


S500
STANDBY
COUNTER = 0
S501
UPWARD ACCELERATION
UPWARD ACCELERATION
S502
UPWARD STOP
STOP
STOP
UPWARD ACCELERATION
UPWARD ACCELERATION
DOWNWARD ACCELERATION
DOWNWARD ACCELERATION
UPDATE COUNTER
UPDATE COUNTER
S503
DOWNWARD ACCELERATION
DOWNWARD ACCELERATION
S504
DOWNWARD STOP
STOP
STOP
DOWNWARD ACCELERATION
0.1 SEC ELAPSED
0.1 SEC ELAPSED
EVENT IF COUNTER = 3

INTERNATIONAL SEARCH REPORT

Ir tional Application No
PCT/JP 02/02298

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 H04N5/262

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G06T H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, PAJ

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5 815 411 A (ELLENBY JOHN ET AL) 29 September 1998 (1998-09-29)	1,2,5,7, 10-12, 14,15
Y		8,9
A	column 2, line 4 -column 3, line 4 column 3, line 53 -column 5, line 48 column 7, line 30 -column 8, line 67 column 15, line 57 - line 66 --- -/--	3,4,6,13

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 July 2002	06/08/2002

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Didierlaurent, P

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/JP 02/02298

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5815411 A	29-09-1998	AU 691508 B2	21-05-1998
		AU 7314694 A	27-03-1995
		CA 2171314 A1	16-03-1995
		CH 689904 A5	14-01-2000
		EP 0722601 A1	24-07-1996
		JP 9505138 T	20-05-1997
		NZ 269318 A	19-12-1997
		US 6031545 A	29-02-2000
		WO 9507526 A1	16-03-1995
		US 5742521 A	21-04-1998
		US 5682332 A	28-10-1997
		US 6307556 B1	23-10-2001
		US 6037936 A	14-03-2000
		US 6064398 A	16-05-2000
		US 6278461 B1	21-08-2001
WO 9703416 A	30-01-1997	EP 0838068 A1	29-04-1998
		JP 11509064 T	03-08-1999
		WO 9703416 A1	30-01-1997
EP 1117074 A	18-07-2001	JP 2001195601 A	19-07-2001
		EP 1117074 A2	18-07-2001